UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/600,797	06/20/2003	Eric D. Brill	MS303968.1 / MSFTP444US	9695

27195	7590	02/28/2007		EXAMINER
AMIN. TUROCY & CALVIN, LLP 24TH FLOOR, NATIONAL CITY CENTER 1900 EAST NINTH STREET CLEVELAND, OH 44114				HICKS, MICHAEL J
			ART UNIT	PAPER NUMBER
			2165	

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	02/28/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

PTOL-90A (Rev. 10/06)

	Application No.	Applicant(s)
	10/600,797	BRILL, ERIC D.
Office Action Summary	Examiner	Art Unit	
	Michael J. Hicks	2165	

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *03 January 2007*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-40,42 and 43* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☐ Claim(s) *1-40, 42-43* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *20 June 2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.

5)☐ Notice of Informal Patent Application

6)☐ Other: _____.

DETAILED ACTION

1.   Claims 1-40 and 42-43 pending.

Claim 41 Canceled.

Response to Arguments

2.   Applicant's arguments with respect to claim 1-43 have been considered but are

moot in view of the new ground(s) of rejection. Note that Applicants arguments are

directed towards the disclosure of Hanson in regard to the limitation of filtering the

search results based on the search entry point. The new art of Lawrence ("Context in Web

Search", IEEE Data Engineering Bulletin, Vol. 23, No. 3, Pgs. 25-32, 2000) is introduced in order to

disclose this limitation, as amended in the current claim listing.

Claim Rejections - 35 USC § 112

3.   The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly
> claiming the subject matter which the applicant regards as his invention.

4.   Claims 2-4 rejected under 35 U.S.C. 112, second paragraph, as being indefinite

for failing to particularly point out and distinctly claim the subject matter which applicant

regards as the invention.

Claim 2 recites the limitation "the criteria" in line 1. There is insufficient

antecedent basis for this limitation in the claim.

Claims 3-4 claim dependencies from Claim 2.

Claim Rejections - 35 USC § 102

5.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

6.   Claims 1-3 rejected under 35 U.S.C. 102(b) as being anticipated by Lawrence.

As per Claim 1, Lawrence discloses a system that refines a general-purpose

search engine (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context*

information, currently in the form of a category of information desired. In addition to providing a keyword

query, users choose a category such as "personal homepages", "research papers", or "general

introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search

engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to

find relevant documents via regular web search engines. The context information is used to select the

search engines to send queries to, to modify queries, and to select the ordering policy." The preceding

text excerpt clearly indicates that a general purpose search engine (e.g. regular web search engine) is

refined using context information.) (Page 1, Column 2, Paragraph 3), comprising: a component that

identifies an entry point that includes a link utilized to access the general-purpose

search engine (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context*

information, currently in the form of a category of information desired. In addition to providing a keyword

query, users choose a category such as "personal homepages", "research papers", or "general

introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search

engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to

find relevant documents via regular web search engines. The context information is used to select the

search engines to send queries to, to modify queries, and to select the ordering policy." The preceding

text excerpt clearly indicates that the entry point, in this case the Inquirus 2 metasearch engine is

identified and comprises a search link that is used to access a general purpose/regular search engine.)

(Page 1, Column 2, Paragraph 3); and a tuning component that filters search query results of

the general-purpose search engine based at least on the entry point through which the

general-purpose search engine was accessed (i.e. *"The Inquirus 2 project at NEC Research*

Institute [29, 30] requests context information, currently in the form of a category of information desired. In

addition to providing a keyword query, users choose a category such as "personal homepages",

"research papers", or "general introductory information". Inquirus 2 is a metasearch engine that operates

as a layer above regular search engines. Inquirus 2 takes a query plus context information, and attempts

to use the context information to find relevant documents via regular web search engines. The context

information is used to select the search engines to send queries to, to modify queries, and to select the

ordering policy." The preceding text excerpt clearly indicates that based on the entry point, in this case

what parameters were given to the entry point, the search query is modified, thus filtering the results of

the general purpose search engine based on the entry point. Note that the general purpose search

engine is accessed by submitting the query, along with the context information to the metasearch engine

(e.g. search link).) (Page 1, Column 2, Paragraph 3).

As per Claim 2, Lawrence discloses the criteria comprising one or more of a

document property, a context parameter, and a configuration (i.e. *"The Inquirus 2 project at*

NEC Research Institute [29, 30] requests context information, currently in the form of a category of

information desired. In addition to providing a keyword query, users choose a category such as "personal

homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine

that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information,

and attempts to use the context information to find relevant documents via regular web search engines.

The context information is used to select the search engines to send queries to, to modify queries, and to

select the ordering policy." The preceding text excerpt clearly indicates that the the criteria may compose

a context parameter (e.g. context information such as a category/parameter selection).) (Page 1, Column

2, Paragraph 3).

Claim 3 holds no patentable weight due to the fact that Claim 2 explicitly states

that the criteria may be something other that a document property. In this case it can be

stated that one of the alternatives is chosen to fulfill the limitation of claim 2, thus

making the details of the document property irrelevant.

Claim Rejections - 35 USC § 103

7.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

8.   Claims 4-40 and 42-43 rejected under 35 U.S.C. 103(a) as being unpatentable

over Lawrence in view of Hansen et al. ("Using navigation data to improve IR functions in the

context of Web search", Proceedings of the tenth international conference on Information and knowledge

management; Atlanta, Georgia, USA; Pages 135-142; ACM Press, 2001 and referred to hereinafter as

Hansen).

As per Claim 4, Lawrence fails to disclose the context parameter comprising one

of a word probability and a probability distribution.

Hansen discloses the context parameter comprising one of a word probability and a probability distribution (i.e. *"The group relation is captures by the triple (q_i, k, w_{ik}), where k denotes a group ID and w_{ik} is the probability that q_i belongs to group k. Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k."* The preceding text excerpt clearly indicates that a relevancy weight/probability distribution is used.) (Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the context parameter comprising one of a word probability and a probability distribution with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 5, Lawrence fails to disclose the tuning component is provided with training data to learn what properties of a document are indicative of the document being relevant to a user executing a search query from the entry point.

Hansen discloses the tuning component provided with training data to learn what properties of a document are indicative of the document being relevant to a user executing a search query from the entry point (i.e. *"To help guide the cluster process, we also introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs from a content hierarchy)."* The preceding text excerpt clearly indicates that training data is provided to help guide the cluster process (e.g. indicate what properties of a document are indicative of that

document being relative to a user).) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraph 1).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the tuning component is provided with training data to learn what properties of a document are indicative of the document being relevant to a user executing a search query from the entry point with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 6, Lawrence fails to disclose the tuning component configured to differentiate between a query result that is relevant to a search query context for a group of users and a query result that is non-relevant to the search query context for the group of users.

Hansen discloses the tuning component configured to differentiate between a query result that is relevant to a search query context for a group of users and a query result that is non-relevant to the search query context for the group of users (i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u$_j$, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u$_j$ is associated with the queries belonging to group k."* The preceding text excerpt clearly indicates that each URL/query result is given a relevance weight associated with the group used for the search. If the URL is relevant, it is assigned a high relevancy score, and if the URL is non-relevant, it is assigned a low relevancy score.) (Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the tuning component configured to differentiate between a query result that is relevant

to a search query context for a group of users and a query result that is non-relevant to

the search query context for the group of users with the motivation of using navigation

(e.g. context) data to improve web searches (Hansen, Abstract).

.As per Claim 7, Lawrence fails to disclose the tuning component employs

statistical analysis in connection with filtering the search query results.

Hansen discloses the tuning component employs statistical analysis in

connection with filtering the search query results (i.e. *"As mentioned above, sets of such triples*

constitute the parameters in a statistical model for the search sessions. These triples can be used by a

search engine to improve page rankings." The preceding text excerpt clearly indicates that statistical

modeling/analysis is used in connection with page ranking/filtering of the search results.) (Page 138,

Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the tuning component employs statistical analysis in connection with filtering the search

query results with the motivation of using navigation (e.g. context) data to improve web

searches (Hansen, Abstract).

As per Claim 8, Lawrence fails to disclose the tuning component generates one

or more context parameters for a received query result, and compares the generated

context parameters with a relevant context parameter and a non-relevant context

parameter to determine whether the query result is relevant.

Hansen discloses the tuning component generates one or more context

parameters for a received query result (i.e. *"The group relation is captures by the triple (q_i, k, w_{ik}),*

where k denotes a group ID and w_{ik} is the probability that q_i belongs to group k. Then, for each group, we

identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a

weight that determines how likely it is that u_j is associated with the queries belonging to group k." The

preceding text excerpt clearly indicates that a relevancy weight/context parameter is generated for a

received query result.) (Page 137, Column 2, Paragraph 4), and then compare the generated

context parameters with a relevant context parameter and a non-relevant context

parameter to determine whether the query result is relevant (i.e. *"The group relation is*

captures by the triple (q_i, k, w_{ik}), where k denotes a group ID and w_{ik} is the probability that q_i belongs to

group k. Then, for each group, we identify a number of relevant URLs. This is described by the triple (k,

u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the

queries belonging to group k." The preceding text excerpt clearly indicates that in order to determine if the

query result is associated with a group (e.g. group k) the relevancy weight/context parameter is

examined. Note that a high relevancy weight constitutes a relevant context parameter and a low

relevancy weight constitutes a non-relevant context parameter. Further note that in order to determine if

the query result is related to the group, the relevancy weight/context parameter for the query result must

be compared to low and high relevancy weights.) (Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the tuning component generates one or more context parameters for a received query

result, and compares the generated context parameters with a relevant context

parameter and a non-relevant context parameter to determine whether the query result

is relevant with the motivation of using navigation (e.g. context) data to improve web

searches (Hansen, Abstract).

As per Claim 9, Lawrence fails to disclose the tuning component further ranks the

query results.

Hansen discloses the tuning component further ranks the query results (i.e. *"Our*

clustering can also be used to modify the rankings of results from a traditional search engine." The

preceding text excerpt clearly indicates that the clustering/filtering is used to improve ranking/rank query

results.) (Page 138, Column 1, Paragraph 3).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the tuning component further ranks the query results with the motivation of using

navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 10, Lawrence fails to disclose the ranking determined by the degree

of relevance of the query result to a relevant data set and a non-relevant data set, the

relevance is determined via one of a similarity measure and a confidence interval.

Hansen discloses the ranking determined by the degree of relevance of the

query result to a relevant data set and a non-relevant data set, the relevance is

determined via one of a similarity measure and a confidence interval (i.e. *"Then, for each*

group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL

and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group

k...As mentioned above, sets of such triples constitute the parameters in a statistical model for the search sessions. These triples can be used by a search engine to improve page rankings." The preceding text excerpt clearly indicates that that ranking is determined by a relevancy score/similarity measure assigned to the URL by comparing it to a group of queries (e.g. group k). Note that not all of the pages in the group that the URL is being compared to are relevant as some may have a low relevancy score in the set, therefore the group (e.g. group k) constitutes both a set of relevant data, and a set of non-relevant data.) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the ranking determined by the degree of relevance of the query result to a relevant data set and a non-relevant data set, the relevance is determined via one of a similarity measure and a confidence interval with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claims 11 and 20, Lawrence fails to disclose the ranking order comprising one of ascending and descending, from the most relevant result to the least relevant result.

Hansen discloses the ranking order comprising one of ascending and descending, from the most relevant result to the least relevant result (i.e. *"Here, we arrange the query groups and the URLs by weight, with the most relevant appearing at the top."* The preceding text excerpt clearly indicates that the results are sorted in descending order, with the most relevant results at the top.) (Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include

the ranking order comprising one of ascending and descending, from the most relevant

result to the least relevant result with the motivation of using navigation (e.g. context)

data to improve web searches (Hansen, Abstract).

As per Claim 12, Lawrence fails to disclose the tuning component configured for

a plurality of entry points associated with one or more groups of users.

Hansen discloses the tuning component configured for a plurality of entry points

associated with one or more groups of users (i.e. "...we present three search sessions (each

initiated by a different user)..." The preceding text excerpt clearly indicates that entry points for multiple

groups of users may be defined.) (Page 137, Column 1, paragraph 3).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the tuning component configured for a plurality of entry points associated with one or

more groups of users with the motivation of using navigation (e.g. context) data to

improve web searches (Hansen, Abstract).

As per Claim 13, Lawrence discloses a system that tunes a general-purpose

search engine (i.e. "The Inquirus 2 project at NEC Research Institute [29, 30] requests context

information, currently in the form of a category of information desired. In addition to providing a keyword

query, users choose a category such as "personal homepages", "research papers", or "general

introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search

engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to

find relevant documents via regular web search engines. The context information is used to select the

search engines to send queries to, to modify queries, and to select the ordering policy." The preceding

text excerpt clearly indicates that a general purpose search engine (e.g. regular web search engine) is

tuned using context information.) (Page 1, Column 2, Paragraph 3), comprising: a filter component

that parses relevant and non-relevant general-purpose search engine content results

depending on the entry point that provides a link employed to traverse to the general-

purpose search engine (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests*

context information, currently in the form of a category of information desired. In addition to providing a

keyword query, users choose a category such as "personal homepages", "research papers", or "general

introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search

engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to

find relevant documents via regular web search engines. The context information is used to select the

search engines to send queries to, to modify queries, and to select the ordering policy." The preceding

text excerpt clearly indicates that based on the entry point, in this case what parameters were given to the

entry point, the search query is modified, thus parsing the results of the general purpose search engine

based on the entry point. Note that the general purpose search engine is accessed by submitting the

query, along with the context information to the metasearch engine (e.g. search link).) (Page 1, Column 2,

Paragraph 3). ·

Lawrence fails to disclose that the parsing is based on training data (i.e. *"To help*

guide the cluster process, we also introduce labeled data from an existing topic hierarchy that contains

over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the training

data, J (which might include URLs from a content hierarchy)." The preceding text excerpt clearly indicates

that training data is provided to help guide the cluster process (e.g. indicate what properties of a

document are indicative of that document being relative to a user).) (Page 137, Column 2, Paragraph 4;

Page 138, Column 1, Paragraph 1), and a ranking component that sorts the filtered results in

accordance with the training data for presentation to a user(i.e. *"To help guide the cluster*

process, we also introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs from a content hierarchy)...These triples can be used by a search engine to improve page rankings." The preceding text excerpt clearly indicates that a ranking component is present which ranks data based on/in accordance with the content hierarchy/training data. Note that Figure 4 illustrated the results being displayed to a user.) (Figure 4; Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraphs 1-2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include that the parsing is based on training data; and a ranking component that sorts the filtered results in accordance with the training data for presentation to a user with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 14, Lawrence discloses the filter component parses the results as a function of one or more of a document property, a context parameter, and a configuration associated with the entry point (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context information, currently in the form of a category of information desired. In addition to providing a keyword query, users choose a category such as "personal homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to find relevant documents via regular web search engines. The context information is used to select the search engines to send queries to, to modify queries, and to select the ordering policy."* The preceding text excerpt clearly indicates that the the criteria may compose a context

parameter (e.g. context information such as a category/parameter selection).) (Page 1, Column 2,

Paragraph 3).

As per Claim 15, Lawrence fails to disclose the filter component trained to

differentiate between a relevant and a non-relevant result via the training data.

Hansen discloses the filter component trained to differentiate between a relevant

and a non-relevant result via the training data (i.e. *"To help guide the cluster process, we also*

introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We let

the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs

from a content hierarchy)...Then, for each group, we identify a number of relevant URLs. This is

described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that

u_j is associated with the queries belonging to group k." The preceding text excerpt clearly indicates that,

using training data, each URL/query result is given a relevance weight associated with the group used for

the search. If the URL is relevant, it is assigned a high relevancy score, and if the URL is non-relevant, it

is assigned a low relevancy score.) (Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the filter component trained to differentiate between a relevant and a non-relevant result.

via the training data with the motivation of using navigation (e.g. context) data to

improve web searches (Hansen, Abstract).

As per Claim 16, Lawrence fails to disclose the training data comprising a set of

relevant data associated with a search context of a user for the entry point and a set of

non-relevant data comprising random data unrelated to the search context of the user
for the entry point.

Hansen discloses the training data comprising a set of relevant data associated
with a search context of a user for the entry point and a set of non-relevant data
comprising random data unrelated to the search context of the user for the entry point
(i.e. *"To help guide the cluster process, we also introduce labeled data from an existing topic hierarchy
that contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in
the training data, J (which might include URLs from a content hierarchy)."* The preceding text excerpt
clearly indicates that a set of training data, the portion of which that is associated with the search context
of the user for the entry point constituting the set of relevant data, and the portion of which that is
irrelevant to the search context of the user for the entry point constituting the set of non-relevant data,
exists and are used to help determine the relevancy of a search result. Note that the set of unrelated
data comes from an existing data hierarchy, and can therefor be considered random.) (Page 137, Column
2, Paragraph 4; Page 138, Column 1, Paragraph 1).

It would have been obvious to one skilled in the art at the time of Applicants
invention to modify the teachings of Lawrence with the teachings of Hansen to include
the training data comprising a set of relevant data associated with a search context of a
user for the entry point and a set of non-relevant data comprising random data
unrelated to the search context of the user for the entry point with the motivation of
using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 17, Lawrence fails to disclose the filter component employs
statistical analysis to determine whether a result is relevant or non-relevant to the entry
point.

Hansen discloses the filter component employs statistical analysis to determine whether a result is relevant or non-relevant to the entry point (i.e. *"As mentioned above, sets of such triples constitute the parameters in a statistical model for the search sessions. These triples can be used by a search engine to improve page rankings."* The preceding text excerpt clearly indicates that statistical modeling/analysis is used in connection with page ranking/filtering of the search results which determine relevancy. Note from above that the triple referred to contain a relevancy score.) (Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the filter component employs statistical analysis to determine whether a result is relevant or non-relevant to the entry point with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 18, Lawrence fails to disclose the ranking component employs a technique to determine the degree of relevance of the query results with respect to a relevant data set and a non-relevant data set.

Hansen discloses the ranking component employs a technique to determine the degree of relevance of the query results with respect to a relevant data set and a non-relevant data set (i.e. *"To help guide the cluster process, we also introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs from a content hierarchy)."* The preceding text excerpt clearly indicates that a set of, the portion of which that is associated with the search context of the user for the entry point constituting the set of relevant data, and the portion of which that is irrelevant to the search context of the user for the entry point constituting the set of non-relevant

data, exists and are used to help determine the relevancy of a query result.) (Page 137, Column 2,
Paragraph 4; Page 138, Column 1, Paragraph 1).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the ranking component employs a technique to determine the degree of relevance of the query results with respect to a relevant data set and a non-relevant data set with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 19, Lawrence fails to disclose the technique comprising one of a similarity measure and a confidence interval.

Hansen discloses the technique comprising one of a similarity measure and a confidence interval (i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k...As mentioned above, sets of such triples constitute the parameters in a statistical model for the search sessions. These triples can be used by a search engine to improve page rankings."* The preceding text excerpt clearly indicates that that ranking/relevancy is determined by a relevancy score/similarity measure assigned to the URL by comparing it to a group of queries (e.g. group k).)(Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the technique comprising one of a similarity measure and a confidence interval with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 21, Lawrence fails to disclose the ranking performed on the

relevant query results, the non-relevant results are discarded.

Hansen discloses the ranking performed on the relevant query results, the non-

relevant results are discarded (i.e. *"For each group, we select the most relevant URLs arranged in a*

display like that in Figure 4. Here, we arrange the query groups and the URLs by weight, with the most

relevant appearing at the top." The preceding text excerpt clearly indicates that only the most relevant

URLs/results are displayed, while the less or non-relevant results are discarded.) (Page 138, Column 1,

Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the ranking performed on the relevant query results, the non-relevant results are

discarded with the motivation of using navigation (e.g. context) data to improve web

searches (Hansen, Abstract).

As per Claim 22, Lawrence discloses a method to filter and rank general-purpose

search engine results associated with an entry point (i.e. *"The Inquirus 2 project at NEC*

Research Institute [29, 30] requests context information, currently in the form of a category of information

desired. In addition to providing a keyword query, users choose a category such as "personal

homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine

that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information,

and attempts to use the context information to find relevant documents via regular web search engines.

The context information is used to select the search engines to send queries to, to modify queries, and to

select the ordering policy." The preceding text excerpt clearly indicates that a general purpose search

engine (e.g. regular web search engine) is refined using context information associated with an entry

point.) (Page 1, Column 2, Paragraph 3), comprising: executing a query search with the

general-purpose search engine accessed through a link associated with the entry point

(i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context information, currently in*

the form of a category of information desired. In addition to providing a keyword query, users choose a

category such as "personal homepages", "research papers", or "general introductory information".

Inquirus 2 is a metasearch engine that operates as a layer above regular search engines. Inquirus 2

takes a query plus context information, and attempts to use the context information to find relevant

documents via regular web search engines. The context information is used to select the search engines

to send queries to, to modify queries, and to select the ordering policy." The preceding text excerpt clearly

indicates that based on the entry point, in this case what parameters were given to the entry point, the

search query is modified, thus filtering the results of the general purpose search engine based on the

entry point. Note that the general purpose search engine is accessed by submitting the query, along with

the context information to the metasearch engine (e.g. search link).) (Page 1, Column 2, Paragraph 3);

filtering the general-purpose search engine results by tuning the general-purpose

search engine based on the entry point employed to access the general purpose search

engine (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context information,*

currently in the form of a category of information desired. In addition to providing a keyword query, users

choose a category such as "personal homepages", "research papers", or "general introductory

information". Inquirus 2 is a metasearch engine that operates as a layer above regular search engines.

Inquirus 2 takes a query plus context information, and attempts to use the context information to find

relevant documents via regular web search engines. The context information is used to select the search

engines to send queries to, to modify queries, and to select the ordering policy." The preceding text

excerpt clearly indicates that based on the entry point, in this case what parameters were given to the

entry point, the search query is modified, thus filtering the results of the general purpose search engine

based on the entry point. Note that the general purpose search engine is accessed by submitting the

query, along with the context information to the metasearch engine (e.g. search link).) (Page 1, Column 2,

Paragraph 3).

Lawrence fails to disclose ranking the general-purpose search engine results.

.Hansen discloses ranking the general-purpose search engine results (i.e. *"These*

triples can be used by a search engine to improve page rankings. When a new user initiates a search,

we present them with a display of query groups most related to their search terms. For each such group,

we select the most relevant URLs arranged in a display like that in Figure 4." The preceding text excerpt

clearly indicates that the general-purpose search engine results are ranked.) (Page 138, Column 1,

Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

ranking the general-purpose search engine results with the motivation of using

navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 23, Lawrence fails to disclose employing a statistical hypothesis to

determine whether a result is relevant or non-relevant to a search context of the entry

point.

Hansen discloses employing a statistical hypothesis to determine whether a

result is relevant or non-relevant to a search context of the entry point (i.e. *"As mentioned*

above, sets of such triples constitute the parameters in a statistical model for the search sessions. These

triples can be used by a search engine to improve page rankings." The preceding text excerpt clearly

indicates that statistical modeling/a statistical hypothesis is used in connection with page ranking of the

search results, which determines relevancy. Note from above that the triple referred to contain a

relevancy score.) (Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include employing a statistical hypothesis to determine whether a result is relevant or non-relevant to a search context of the entry point with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 24, Lawrence fails to disclose the statistical hypothesis employing a threshold in connection with a probability distribution for relevant data and a probability distribution for non-relevant data, respective word probabilities are generated for the search query results and compared to the threshold, the probability distribution for relevant data and the probability distribution for non-relevant data to determine whether the results are relevant or non-relevant.

Hansen discloses the statistical hypothesis employing a threshold in connection with a probability distribution for relevant data and a probability distribution for non-relevant data (i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k...As mentioned above, sets of such triples constitute the parameters in a statistical model for the search session...For each such group, we select the most relevant URLs..."* The preceding text excerpt clearly indicates that a threshold to determine which are the most relevant URLs exists and that a relevancy weight/probability distribution of a document is calculated for each group, some of which are relevant to the search, and some of which are not.) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraphs 1-2), wherein respective word probabilities are generated for the search query results and compared to the threshold

(i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k...For each such group, we select the most relevant URLs..."* The preceding text excerpt clearly indicates that in order for a URL/result to be considered most relevant it's relevancy weight/probability distribution must be first calculated, then compared to the threshold.) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraphs 1-2), the probability distribution for relevant data and the probability distribution for non-relevant data to determine whether the results are relevant or non-relevant (i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k."* The preceding text excerpt clearly indicates that the relevancy score/probability distribution is used to determine which results are relevant and which results are non-relevant.) (Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the statistical hypothesis employing a threshold in connection with a probability distribution for relevant data and a probability distribution for non-relevant data, respective word probabilities are generated for the search query results and compared to the threshold, the probability distribution for relevant data and the probability distribution for non-relevant data to determine whether the results are relevant or non-relevant with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 25, Lawrence fails to disclose the threshold employed to bias the decision based on one of a result being deemed non-relevant when the result is relevant and a result being deemed relevant when the result is non-relevant.

Hansen discloses the threshold employed to bias the decision to mitigate one of a result being deemed non-relevant when the result is relevant and a result being deemed relevant when the result is non-relevant (i.e. *"We capture the interesting part of the search path in a search session, which is the user's query together with the URLs of the Web sites they visit in response to their query...Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k...As mentioned above, sets of such triples constitute the parameters in a statistical model for the search sessions."* The preceding text excerpt clearly indicates that because the results are collected over many search sessions, and are collected from which URLs users visit, if one users visits a non-relevant URL in response to a search, this URL will only be determined as relevant in response to one search session. Unless the data from the collected search sessions also indicate the non-relevant URL, its relevancy weight will not be greatly affected, therefore mitigating/lessening the impact. Also not that, in much the same way, relevant URLs relevancy weights will not be greatly affected as a result of being deemed non-relevant in one particular search session. The data from many collected search sessions is needed to markedly influence a relevancy weight.) (Page 135, Column 2, Paragraph 2; Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the threshold employed to bias the decision based on one of a result being deemed non-relevant when the result is relevant and a result being deemed relevant when the

result is non-relevant with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 26, Lawrence fails to disclose further employing a probability distribution analysis or machine learning in connection with the filtering and ranking, wherein suitable probability distributions include a Bernoulli, a binomial, a Pascal, a Poisson, an arcsine, a beta, a Cauchy, a chi-square with N degrees of freedom, an Erlang, a uniform, an exponential, a gamma, a Gaussian-univariate, a Gaussian-bivariate, a Laplace, a log-normal, a rice, a Weibull and a Rayleigh distribution, and the machine learning can classify based on one or more of a word occurrence, a distribution, a page layout, an inlink, and an outlink.

Hansen discloses employing a probability distribution analysis or machine learning in connection with the filtering and ranking, wherein suitable probability distributions include a Bernoulli, a binomial, a Pascal, a Poisson, an arcsine, a beta, a Cauchy, a chi-square with N degrees of freedom, an Erlang, a uniform, an exponential, a gamma, a Gaussian-univariate, a Gaussian-bivariate, a Laplace, a log-normal, a rice, a Weibull and a Rayleigh distribution (i.e. *"The same kind of Poisson structure usd for the collection of URLs in a search session is applied to the query terms."* The preceding text excerpt clearly indicates that a poisson distribution is used.) (Page 11, Column 1, Paragraph 6), and the machine learning can classify based on one or more of a word occurrence, a distribution, a page layout, an inlink, and an outlink (i.e. *"A natural prior for our coefficients λ_{ij} (the relevance weights) is a Gamma distribution."* The preceding text excerpt clearly indicates a distribution is used to classify the relevancies.) (Page 141, Column 1, Paragraph 5).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

further employing a probability distribution analysis or machine learning in connection

with the filtering and ranking, wherein suitable probability distributions include a

Bernoulli, a binomial, a Pascal, a Poisson, an arcsine, a beta, a Cauchy, a chi-square

with N degrees of freedom, an Erlang, a uniform, an exponential, a gamma, a

Gaussian-univariate, a Gaussian-bivariate, a Laplace, a log-normal, a rice, a Weibull

and a Rayleigh distribution, and the machine learning can classify based on one or

more of a word occurrence, a distribution, a page layout, an inlink, and an outlink with

the motivation of using navigation (e.g. context) data to improve web searches (Hansen,

Abstract).

As per Claim 27, Lawrence fails to disclose employing a statistical analysis to

rank search query results.

Hansen discloses employing a statistical analysis to rank search query results

(i.e. *"As mentioned above, sets of such triples constitute the parameters in a statistical model for the*

search sessions. These triples can be used by a search engine to improve page rankings." The

preceding text excerpt clearly indicates that statistical modeling/analysis is used in connection with page

ranking of the search results.) (Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

employing a statistical analysis to rank search query results with the motivation of using

navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 28, Lawrence fails to disclose the ranking comprising one of generating word probabilities and employing a confidence interval to determine relevance, and generating a similarity measure comprising one of a cosine distance, the Jaccard coefficient, an entropy-based measure, a divergence measure and/or a relative separation measure to determine similarity.

Hansen discloses the ranking comprising one of generating word probabilities and employing a confidence interval to determine relevance, and generating a similarity measure comprising one of a cosine distance, the Jaccard coefficient, an entropy-based measure, a divergence measure and/or a relative separation measure to determine similarity (i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k."* The preceding text excerpt clearly indicates that a weight is given to each site to identify how closely it is related to a group k, and which is used in ranking. Note that the weights are based on the search terms, as in Figures 3 and 4, which therefore indicates that word probabilities are used to assign the weights and ranks.) (Figures 3, 4; Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the ranking comprising one of generating word probabilities and employing a confidence interval to determine relevance, and generating a similarity measure comprising one of a cosine distance, the Jaccard coefficient, an entropy-based measure, a divergence measure and/or a relative separation measure to determine similarity with the

motivation of using navigation (e.g. context) data to improve web searches (Hansen,

Abstract).

As per Claim 29, Lawrence discloses a method to customize a general-purpose

search engine to improve context search query results (i.e. *"The Inquirus 2 project at NEC*

Research Institute [29, 30] requests context information, currently in the form of a category of information

desired. In addition to providing a keyword query, users choose a category such as "personal

homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine

that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information,

and attempts to use the context information to find relevant documents via regular web search engines.

The context information is used to select the search engines to send queries to, to modify queries, and to

select the ordering policy." The preceding text excerpt clearly indicates that a general purpose search

engine (e.g. regular web search engine) is refined using context information.) (Page 1, Column 2,

Paragraph 3), comprising: tuning a general-purpose search engine for an entry point by

employing a method (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context*

information, currently in the form of a category of information desired. In addition to providing a keyword

query, users choose a category such as "personal homepages", "research papers", or "general

introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search

engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to

find relevant documents via regular web search engines. The context information is used to select the

search engines to send queries to, to modify queries, and to select the ordering policy." The preceding

text excerpt clearly indicates that a general purpose search engine (e.g. regular web search engine) is

refined using context information.) (Page 1, Column 2, Paragraph 3) further comprising: the entry

point provides a link employed to access the general-purpose search engine (i.e. *"The*

Inquirus 2 project at NEC Research Institute [29, 30] requests context information, currently in the form of

a category of information desired. In addition to providing a keyword query, users choose a category such as "personal homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to find relevant documents via regular web search engines. The context information is used to select the search engines to send queries to, to modify queries, and to select the ordering policy." The preceding text excerpt clearly indicates that the entry point, in this case the Inquirus 2 metasearch engine is identified and comprises a search link that is used to access a general purpose/regular search engine.) (Page 1, Column 2, Paragraph 3).

Lawrence fails to disclose providing a set of relevant data to train a component to discern query results relevant to a search context of a user employing the entry point, providing a set of non-relevant data to train the component to discern query results unrelated to the search context, the set of relevant data and the set of non-relevant data are manually provided; and determining whether a query result is relevant to the search context based on the set of relevant data and the set of non-relevant data.

Hansen discloses providing a set of relevant data to train a component to discern query results relevant to a search context of a user employing the entry point (i.e. *"To help guide the cluster process, we also introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs from a content hierarchy)."* The preceding text excerpt clearly indicates a set of data, the portion of which is related to the query constituting the set of relevant data, which guides the cluster process/is used to discern query results relevant to a search context of a user employing the entry point. Note that URLs relevance will be measured by a relevancy score which pertains to how relevant the data is to the groups in the training set. A high relevancy score indicates a relevant URL.) (Page 137, Column 2, Paragraph 3); providing a set of non-relevant data to train the component to discern query results unrelated to the search context (i.e. *"To help guide*

the cluster process, we also introduce labeled data from an existing topic hierarchy that contains over 1.2

million Web sites...We let the index j range from 1 to the number of URLs seen in the training data, J

(which might include URLs from a content hierarchy)." The preceding text excerpt clearly indicates a set

of data, the portion of which is unrelated to the query constituting the set of non-relevant data, which

guides the cluster process/is used to discern query results non-relevant to a search context of a user

employing the entry point. Note that URLs relevance will be measured by a relevancy score which

pertains to how relevant the data is to the groups in the training set. A low relevancy score indicates an

unrelated URL.) (Page 137, Column 2, Paragraph 3), the set of relevant data and the set of non-

relevant data are manually provided; and determining whether a query result is relevant

to the search context based on the set of non-relevant data (i.e. *"To help guide the cluster*

process, we also introduce labeled data from an existing topic hierarchy that contains over 1.2 million

Web sites...We let the index j range from 1 to the number of URLs seen in the training data, J (which

might include URLs from a content hierarchy)." The preceding text excerpt clearly indicates that the set of

relevant data and the set of non-relevant data are manually provided and employed to guide the

clustering process/determine whether a query result is relevant to a search context.) (Page 137, Column

2, Paragraph 3).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

providing a set of relevant data to train a component to discern query results relevant to

a search context of a user employing the entry point, providing a set of non-relevant

data to train the component to discern query results unrelated to the search context, the

set of relevant data and the set of non-relevant data are manually provided; and

determining whether a query result is relevant to the search context based on the set of

relevant data and the set of non-relevant data with the motivation of using navigation

(e.g. context) data to improve web searches (Hansen, Abstract).

As per Claims 30 and 35, Lawrence fails to disclose the set of relevant data

comprising data associated with the search context of the user for the entry point.

Hansen discloses the set of relevant data comprising data associated with the

search context of the user for the entry point (i.e. *"To help guide the cluster process, we also*

introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We let

the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs

from a content hierarchy)." The preceding text excerpt clearly indicates that a set of data, the portion of

which that is associated with the search context of the user for the entry point constituting the set of

relevant data, exists.) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraph 1).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the set of relevant data comprising data associated with the search context of the user

for the entry point with the motivation of using navigation (e.g. context) data to improve

web searches (Hansen, Abstract).

As per Claims 31 and 36, Lawrence fails to disclose the set of non-relevant data

comprising random data unrelated to the search context of the user for the entry point.

Hansen discloses the set of non-relevant data comprising random data unrelated

to the search context of the user for the entry point (i.e. *"To help guide the cluster process, we*

also introduce labeled data from an existing topic hierarchy that contains over 1.2 million Web sites...We

let the index j range from 1 to the number of URLs seen in the training data, J (which might include URLs from a content hierarchy)." The preceding text excerpt clearly indicates that a set of data, the portion of which that is unrelated with the search context of the user for the entry point constituting the set of non-relevant data, exists.) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraph 1).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include the set of non-relevant data comprising random data unrelated to the search context of the user for the entry point with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claims 32 and 37, Lawrence fails to disclose providing information to associate respective query results with the entry point.

Hansen discloses providing information to associate respective query results with the entry point (i.e. *"Then, for each group, we identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a weight that determines how likely it is that u_j is associated with the queries belonging to group k...For each such group, we select the most relevant URLs..."* The preceding text excerpt clearly indicates that the relevancy weight is used to associate query results with the relevant group/the entry point) (Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include providing information to associate respective query results with the entry point with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claims 33 and 38, Lawrence fails to disclose the set of relevant data and

the set of non-relevant data employed to train the component to learn the features that

differentiate relevant data from non-relevant data.

Hansen discloses the set of relevant data and the set of non-relevant data

employed to train the component to learn the features that differentiate relevant data

from non-relevant data (i.e. *"To help guide the cluster process, we also introduce labeled data from*

an existing topic hierarchy that contains over 1.2 million Web sites...We let the index j range from 1 to the

number of URLs seen in the training data, J (which might include URLs from a content hierarchy)." The

preceding text excerpt clearly indicates that the set of relevant data and the set of non-relevant data are

used as a set of training data to identify relevant data.) (Page 137, Column 2, Paragraph 4; Page 138,

Column 1, Paragraph 1).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

the set of relevant data and the set of non-relevant data employed to train the

component to learn the features that differentiate relevant data from non-relevant data

with the motivation of using navigation (e.g. context) data to improve web searches

(Hansen, Abstract).

As per Claim 34, Lawrence discloses a method to automatically customize a

general-purpose search engine for an entry point (i.e. *"The Inquirus 2 project at NEC Research*

Institute [29, 30] requests context information, currently in the form of a category of information desired. In

addition to providing a keyword query, users choose a category such as "personal homepages",

"research papers", or "general introductory information". Inquirus 2 is a metasearch engine that operates

'as a layer above regular search engines. *Inquirus 2 takes a query plus context information, and attempts to use the context information to find relevant documents via regular web search engines. The context information is used to select the search engines to send queries to, to modify queries, and to select the ordering policy."* The preceding text excerpt clearly indicates that a general purpose search engine (e.g. regular web search engine) is refined using context information.) (Page 1, Column 2, Paragraph 3),

comprising: identifying the entry point (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context information, currently in the form of a category of information desired. In addition to providing a keyword query, users choose a category such as "personal homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to find relevant documents via regular web search engines. The context information is used to select the search engines to send queries to, to modify queries, and to select the ordering policy."* The preceding text excerpt clearly indicates that the entry point, in this case the Inquirus 2 metasearch engine is identified and comprises a search link that is used to access a general purpose/regular search engine.) (Page 1, Column 2, Paragraph 3); executing a query search via the entry point that includes a link employed to route to the general-purpose search engine (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30] requests context information, currently in the form of a category of information desired. In addition to providing a keyword query, users choose a category such as "personal homepages", "research papers", or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context information to find relevant documents via regular web search engines. The context information is used to select the search engines to send queries to, to modify queries, and to select the ordering policy."* The preceding text excerpt clearly indicates that the entry point, in this case the Inquirus 2 metasearch engine is identified and comprises a search link that is used to access a general purpose/regular search engine.) (Page 1, Column 2, Paragraph 3).

Lawrence fails to disclose recording a query result selected by a user employing

the entry point as relevant; recording a higher ranked query results as non-relevant,

when a lower ranked result is selected by the user, and providing the recorded results to

automatically train the filter for the entry point, in order to discriminate between results

relevant to a search context of the user for the entry point and results non-relevant to

the search context.

Hansen discloses recording a query result selected by a user as relevant (i.e. *"We*

capture the interesting part of the search path in a search session, which is a user's query together with

the URLs of the Web pages they visit in response to their query...by combining search sessions with

queries in a given group, we can better identify relevant URLs." The preceding text excerpt clearly

indicates that URLs/query results selected by a user are used as data to identify relevant URLs, therefore

some of them will be marked as relevant.) (Page 135, Column 2, Paragraph 2; Page 137, Column 1,

Paragraph 3); recording a higher ranked query results, wherein a lower ranked result is

selected by the user, as non-relevant (i.e. *"...PageRank is based on the amount of time a "random*

surfer" would spend on each page." The preceding text excerpt clearly indicates that if the browser

spends time on a given web page, it increases the relevancy of that web page (e.g. marks it as relevant)

and decreases the relevancy of other pages, which may be considered more relevant (e.g. marks a

relevant page as non-relevant).) (Page 137, Column 2, Paragraph 2); and providing the recorded

results to automatically train the filter to discriminate between results relevant to a

search context of the user for the entry point and results non-relevant to the search

context (i.e. *"We capture the interesting part of the search path in a search session, which is a user's*

query together with the URLs of the Web pages they visit in response to their query...Implicit in our

approach is a form of query clustering that combines similar search terms on the basis of Web pages

visited during a search session. These clusters are then used to improve the display of search engine

results." The preceding text excerpt clearly indicates that previous query results and selections are recorded and added to the set of data used to discriminate between relevant results and non-relevant results.) (Page 135, Column 2, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include recording a query result selected by a user employing the entry point as relevant; recording a higher ranked query results as non-relevant, when a lower ranked result is selected by the user, and providing the recorded results to automatically train the filter for the entry point, in order to discriminate between results relevant to a search context of the user for the entry point and results non-relevant to the search context with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 39, Lawrence fails to disclose the query results selected via a click thru technique employing a mouse to select a link associated with the query result by clicking on the link.

Hansen discloses the query results selected via a click thru technique employing a mouse to select a link associated with the query result by clicking on the link (i.e. "...users click through data to discover disjoint sets of similar URLs" The preceding text excerpt clearly indicates results and URLs are viewed using a click-through technique incorporating a mouse.) (Page 141, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include

the query results selected via a click thru technique employing a mouse to select a link

associated with the query result by clicking on the link with the motivation of using

navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 40, Lawrence fails to disclose generating a word probability

distribution for the relevant recorded results and a word probability distribution for the

non-relevant recorded results.

Hansen discloses generating a word probability distribution for the relevant

recorded results and a word probability distribution for the non-relevant recorded results

(i.e. *"To help guide the cluster process, we also introduce labeled data from an existing topic hierarchy*

that contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in

the training data, J (which might include URLs from a content hierarchy)...Then, for each group, we

identify a number of relevant URLs. This is described by the triple (k, u_j, λ_{kj}) where uj is a URL and λ_{kj} is a

weight that determines how likely it is that u_j is associated with the queries belonging to group k." The

preceding text excerpt clearly indicates that a weight is given to each site to identify how closely it is

related to a group k, and which is used in ranking. Note that the weights are based on the search terms,

as in Figures 3 and 4, which therefore indicates that word probabilities are used to assign the weights and

ranks. Note that the weights are generated for each group and that the data set is comprised of a portion

that is associated with the search context of the user for the entry point and constitutes the set of relevant

data, and a portion that is irrelevant to the search context of the user for the entry point and constitutes

the set of non-relevant data.) (Figures 3, 4; Page 137, Column 2, Paragraph 4).

It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include

generating a word probability distribution for the relevant recorded results and a word

probability distribution for the non-relevant recorded results with the motivation of using

navigation (e.g. context) data to improve web searches (Hansen, Abstract).

As per Claim 42, Lawrence discloses a computer readable medium storing

computer executable components that tunes a general-purpose search engine to

improve context search query results (i.e. *"The Inquirus 2 project at NEC Research Institute [29,*

30] requests context information, currently in the form of a category of information desired. In addition to

providing a keyword query, users choose a category such as "personal homepages", "research papers",

or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above

regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context

information to find relevant documents via regular web search engines. The context information is used to

select the search engines to send queries to, to modify queries, and to select the ordering policy." The

preceding text excerpt clearly indicates that a general purpose search engine (e.g. regular web search

engine) is refined using context information.) (Page 1, Column 2, Paragraph 3), comprising: filtering

search results based on the entry point that provides a link utilized to arrive at the

general-purpose search engine (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30]*

requests context information, currently in the form of a category of information desired. In addition to

providing a keyword query, users choose a category such as "personal homepages", "research papers",

or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above

regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context

information to find relevant documents via regular web search engines. The context information is used to

select the search engines to send queries to, to modify queries, and to select the ordering policy." The

preceding text excerpt clearly indicates that the entry point, in this case the Inquirus 2 metasearch engine

is identified and comprises a search link that is used to access a general purpose/regular search engine.)

(Page 1, Column 2, Paragraph 3).

Lawrence fails to disclose a component that filters the general-purpose search

engine results based on training data sets associated with the search context of a user

and a component that ranks the general-purpose search engine results according to the

similarity of the search engine results to the training data sets.

Hansen discloses a component that filters the general-purpose search engine

results based on training data sets associated with the search context of a user (i.e. *"To*

help guide the cluster process, we also introduce labeled data from an existing topic hierarchy that

contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the

training data, J (which might include URLs from a content hierarchy)...When a new user initiates a

search, we present them with a display of query groups most related to their search terms. For each such

group, we select the most relevant URLs arranged in a display like that in Figure 4." The preceding text

excerpt clearly indicates training data associated with the search context of a user is used to filter general

purpose search engine results.) (Figure 4; Page 137, Column 2, Paragraph 4; Page 138, Column 1,

Paragraphs 1-2) and a component that ranks the general-purpose search engine results

according to the similarity of the search engine results to the training data sets (i.e. *"To*

help guide the cluster process, we also introduce labeled data from an existing topic hierarchy that

contains over 1.2 million Web sites...We let the index j range from 1 to the number of URLs seen in the

training data, J (which might include URLs from a content hierarchy)...These triples can be used by a

search engine to improve page rankings." The preceding text excerpt clearly indicates that a ranking

component is present which ranks data based on/in accordance with similarity to the content

hierarchy/training data.) (Page 137, Column 2, Paragraph 4; Page 138, Column 1, Paragraphs 1-2).

.It would have been obvious to one skilled in the art at the time of Applicants

invention to modify the teachings of Lawrence with the teachings of Hansen to include a

component that filters the general-purpose search engine results based on training data

sets associated with the search context of a user and a component that ranks the

general-purpose search engine results according to the similarity of the search engine

results to the training data sets with the motivation of using navigation (e.g. context)

data to improve web searches (Hansen, Abstract).

As per Claim 43, Lawrence discloses a system that filters and ranks general-

purpose search engine results (i.e. *"The Inquirus 2 project at NEC Research Institute [29, 30]*

requests context information, currently in the form of a category of information desired. In addition to

providing a keyword query, users choose a category such as "personal homepages", "research papers",

or "general introductory information". Inquirus 2 is a metasearch engine that operates as a layer above

regular search engines. Inquirus 2 takes a query plus context information, and attempts to use the context

information to find relevant documents via regular web search engines. The context information is used to

select the search engines to send queries to, to modify queries, and to select the ordering policy." The

preceding text excerpt clearly indicates that a general purpose search engine (e.g. regular web search

engine) is refined using context information.) (Page 1, Column 2, Paragraph 3), comprising: means

for filtering general-purpose search engine results to determine whether a query result

is relevant to a search context of a group of users depending on the entry point that

includes a link employed to navigate to the general-purpose search engine (i.e. *"The*

Inquirus 2 project at NEC Research Institute [29, 30] requests context information, currently in the form of

a category of information desired. In addition to providing a keyword query, users choose a category such

as "personal homepages", "research papers", or "general introductory information". Inquirus 2 is a

metasearch engine that operates as a layer above regular search engines. Inquirus 2 takes a query plus

context information, and attempts to use the context information to find relevant documents via regular

web search engines. The context information is used to select the search engines to send queries to, to

modify queries, and to select the ordering policy." The preceding text excerpt clearly indicates that based

on the entry point, in this case what parameters were given to the entry point, the search query is modified, thus filtering the results of the general purpose search engine based on the entry point. Note that the general purpose search engine is accessed by submitting the query, along with the context information to the metasearch engine (e.g. search link).) (Page 1, Column 2, Paragraph 3).

Lawrence fails to disclose means for ranking the general-purpose search engine results based on a relevance of the general-purpose search engine results to the search context of the group of users and the entry point.

Hansen discloses means for ranking the general-purpose search engine results based on a relevance of the general-purpose search engine results to the search context of the group of users (i.e. *"These triples can be used by a search engine to improve page rankings. When a new user initiates a search, we present them with a display of query groups most related to their search terms. For each such group, we select the most relevant URLs arranged in a display like that in Figure 4."* The preceding text excerpt clearly indicates that the general-purpose search engine results are ranked based on relevance to the search terms/context of the users.) (Page 138, Column 1, Paragraph 2).

It would have been obvious to one skilled in the art at the time of Applicants invention to modify the teachings of Lawrence with the teachings of Hansen to include means for ranking the general-purpose search engine results based on a relevance of the general-purpose search engine results to the search context of the group of users with the motivation of using navigation (e.g. context) data to improve web searches (Hansen, Abstract).

Points of Contact

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael J. Hicks whose telephone number is (571) 272-2670. The examiner can normally be reached on Monday - Friday 10:00a - 7:00p.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jeffrey Gaffin can be reached on (571) 272-4146. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Michael J Hicks
Art Unit 2165
Phone: (571) 272-2670
Fax: (571) 273-2670

JEFFREY GAFFIN
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100